



Le projet EU DataGrid et la participation d'ATLAS

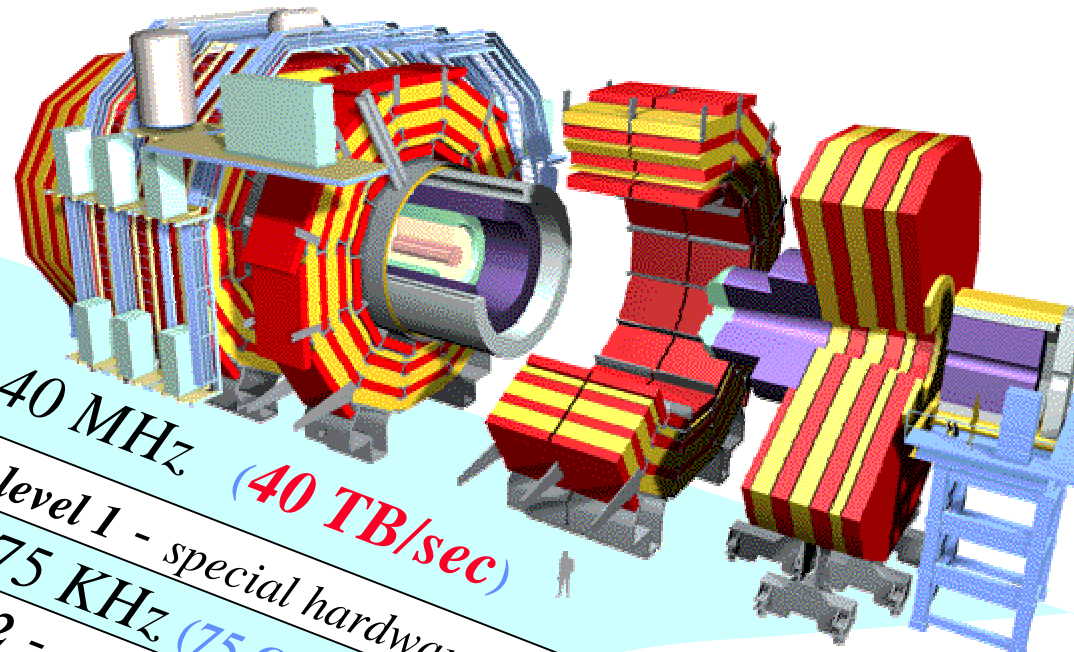
La symbolique de GRID



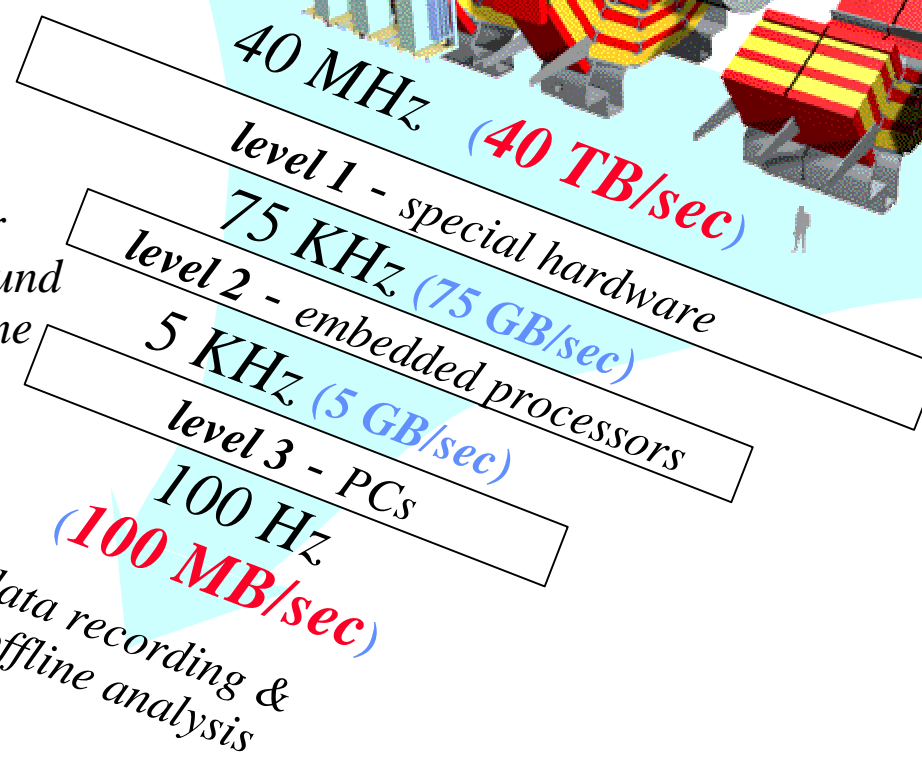
- Calcul distribué inévitablement illimité
- Accès transparent aux multipetabytes de bases de données distribuées
- Facile à mettre en place
- La complexité de la structure est cachée
- Analogie avec le réseau électrique



*Une des 4
expériences sur
le LHC*



online system
*multi-level trigger
filter out background
reduce data volume*

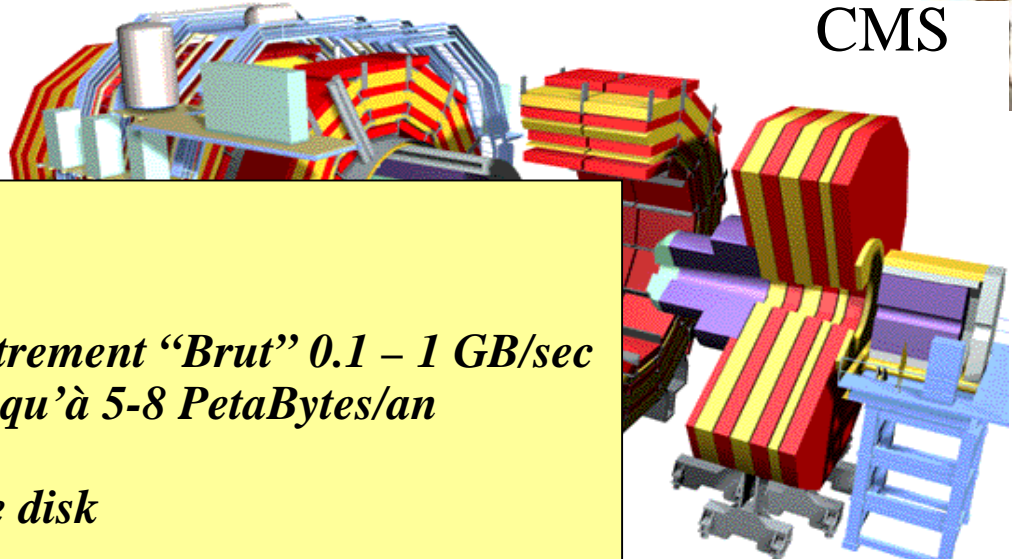
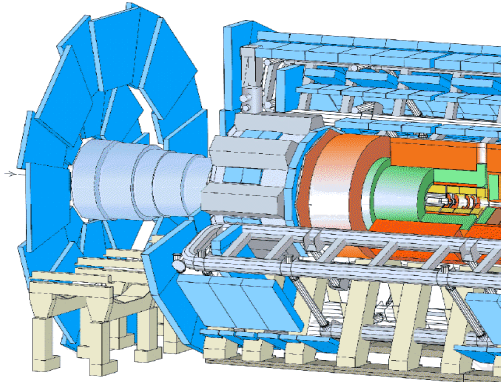


Les détecteurs du LHC



CMS

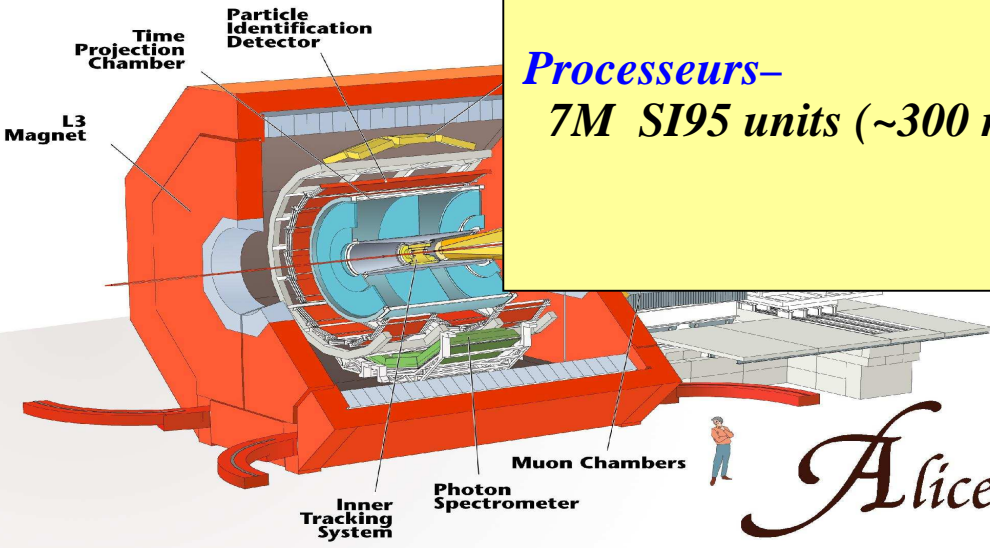
ATLAS



Stockage –
 Taux d'enregistrement "Brut" 0.1 – 1 GB/sec
 accumulant jusqu'à 5-8 PetaBytes/an

10 PetaBytes de disk

Processeurs–
 7M SI95 units (~300 million MIPS)



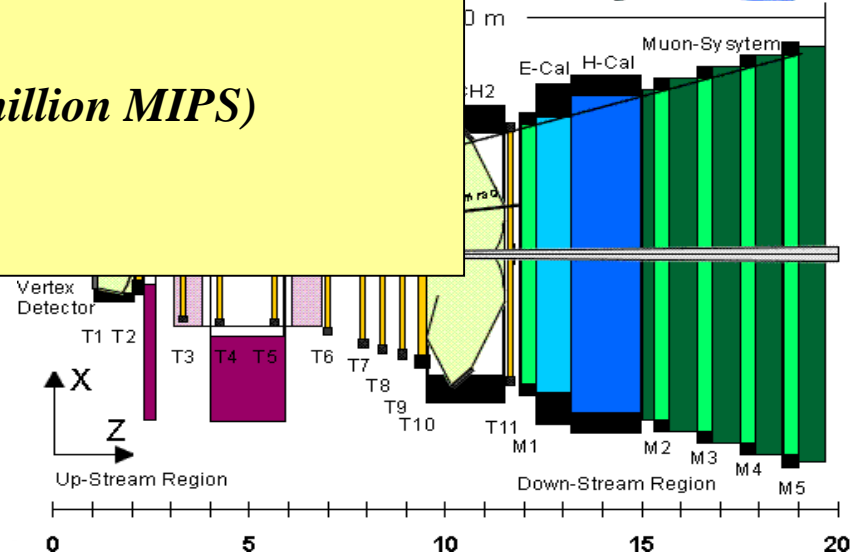
Time Projection Chamber
 Particle Identification Detector

L3 Magnet

Inner Tracking System
 Photon Spectrometer

Muon Chambers

Alice



Le réseau CERN dans le monde



Europe: 267 instituts, 4603 utilisateurs

Ailleurs: 208 instituts, 1632 utilisateurs 5

Le Problème



- Quantité Vaste de données
- Demandes de calculs énormes
- Des chercheurs partout dans le monde

La physique est A U S S I un Challenge !

La solution 1 : distribuer le travail et les données le cas d'ATLAS



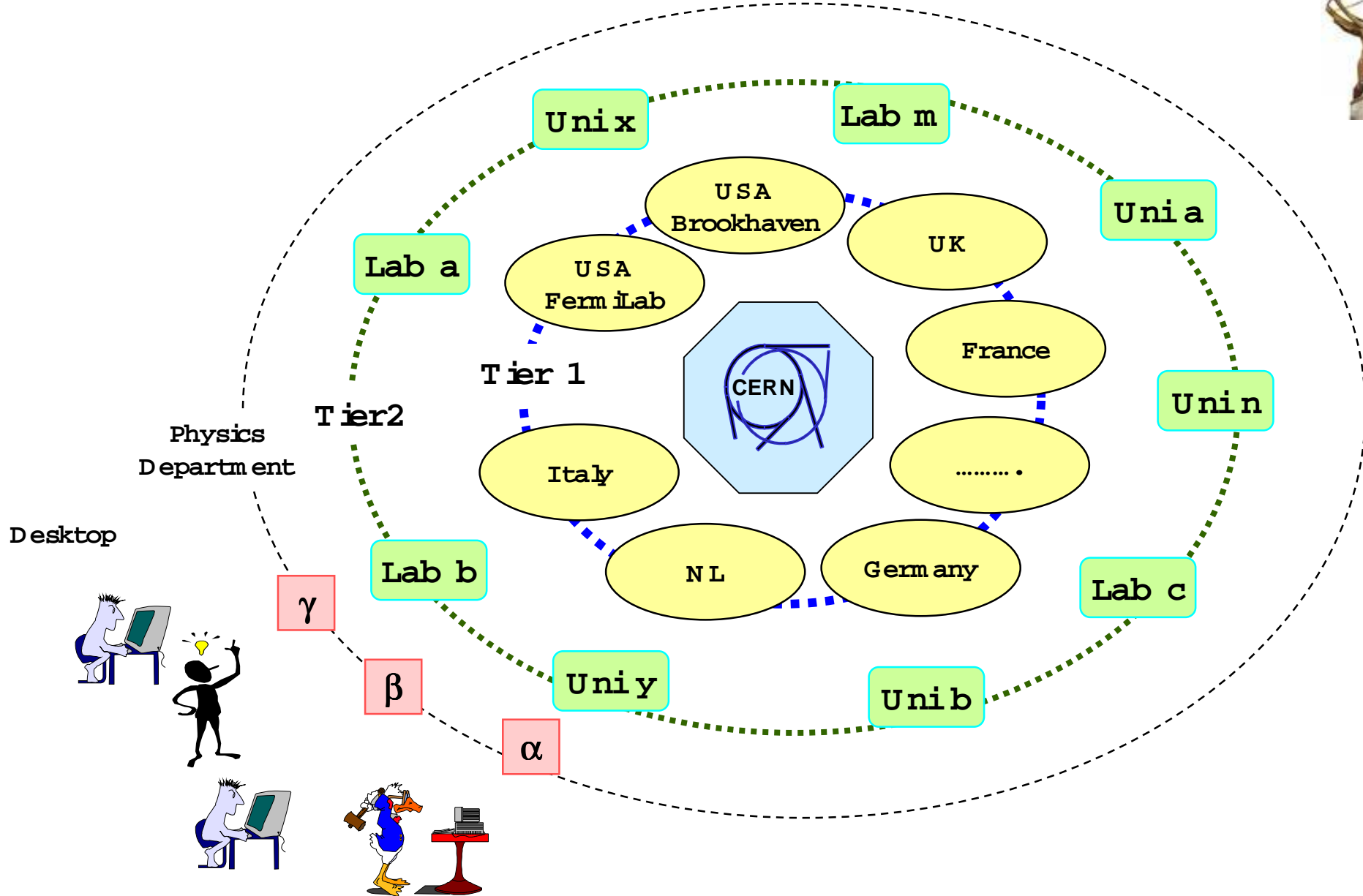
- Mise en place des Data-Challenges: Simulation du détecteur, reconstruction etc ...
 - DC0: 10^5 collisions p-p → été 2002
 - DC1:
 - Phase 1: 10^7 collisions p-p et $3 \cdot 10^7$ single → 25 TB données produites → décembre 2002
 - Phase 2: Données Phase 1 à reconstruire
 - DC2, DC3 ... → 2003 à 2005
- Participation de 150 universités et laboratoires: 34 pays
 - DC1: 15 pays dont la France
- Délocalisation des données, distribution du software sous forme de RPM s, partage des taches
- Conséquences: beaucoup de burdeur dans les transferts de données (bbftp, ftp, ...), des procédures de validation, des publications de l'information (BookKeeping) etc..

La solution 2 : Utiliser la technologie Grid qu'on on peut



- **En principe, Vous soumettez une tâche et Grid**
 - Trouve l'endroit idéal pour que cette tâche fonctionne
 - Organise des accès efficaces à vos données
 - **lecture, migration, replication**
 - "Discute" avec les autorités de certification des différents sites que vous accédez
 - Interface au site local pour les ressources, les mécanismes d'allocations, les règles d'usage etc.. .
 - Fais fonctionner les tâches soumises
 - Monitore le progrès
 - .. et.. vous contacte pour vous prévenir si votre tâche est finie

Le modèle LHC Computing



Programme EDG

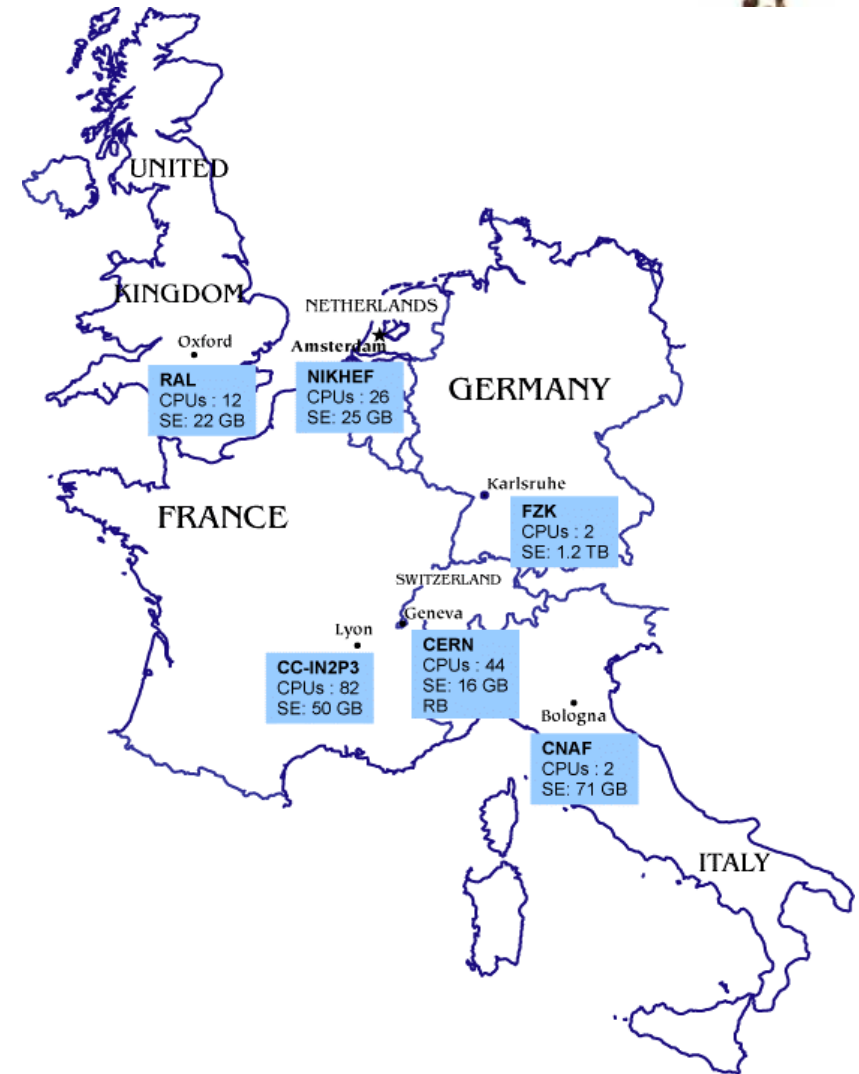


- Mandat: 1er Janvier 2001, fin 2003
 - But: développement d'un "Grid middleware" pour la Physique des Hautes Energies, l'observation de la terre et les applications en Biologie
- Le coeur du "testbed" consiste en un site central, le CERN et d'autres facilités en Europe de l'ouest.
- Nous avons atteint un niveau de stabilité "suffisant" pour faire des tests de production à grande échelle.

Le Testbed EDG et l'implication d'ATLAS



- EDG est impliqué dans la création d'un testbed stable pour une utilisation d'applications
 - démarrage en août 2002
 - ...et coïncide avec ATLAS DC1
 - La priorité a donc été donnée à ATLAS
- Tous les Sites utilisent les outils EDG (basés sur RedHat 6.2)
- L'élément Central: Resource Broker (RB), distribue les jobs entre les différentes ressources
 - Maintenant: , 3 RB au CERN , 2 à LNFN
 - Futur, 1 RB par "Virtual Organization" (VO)





Les services offerts par EDG

▪ *Workload Management*

- *spécification des options de jobs*
- *input data - (données en entrée) conditionne les soumissions de job*
- *"les meilleures" ressources sont choisies*

(si pas de données en entrée, ou d'autres ressources équivalentes ..)

▪ *Information System* : <http://www.nordugrid.org/monitor/edgstat/>?

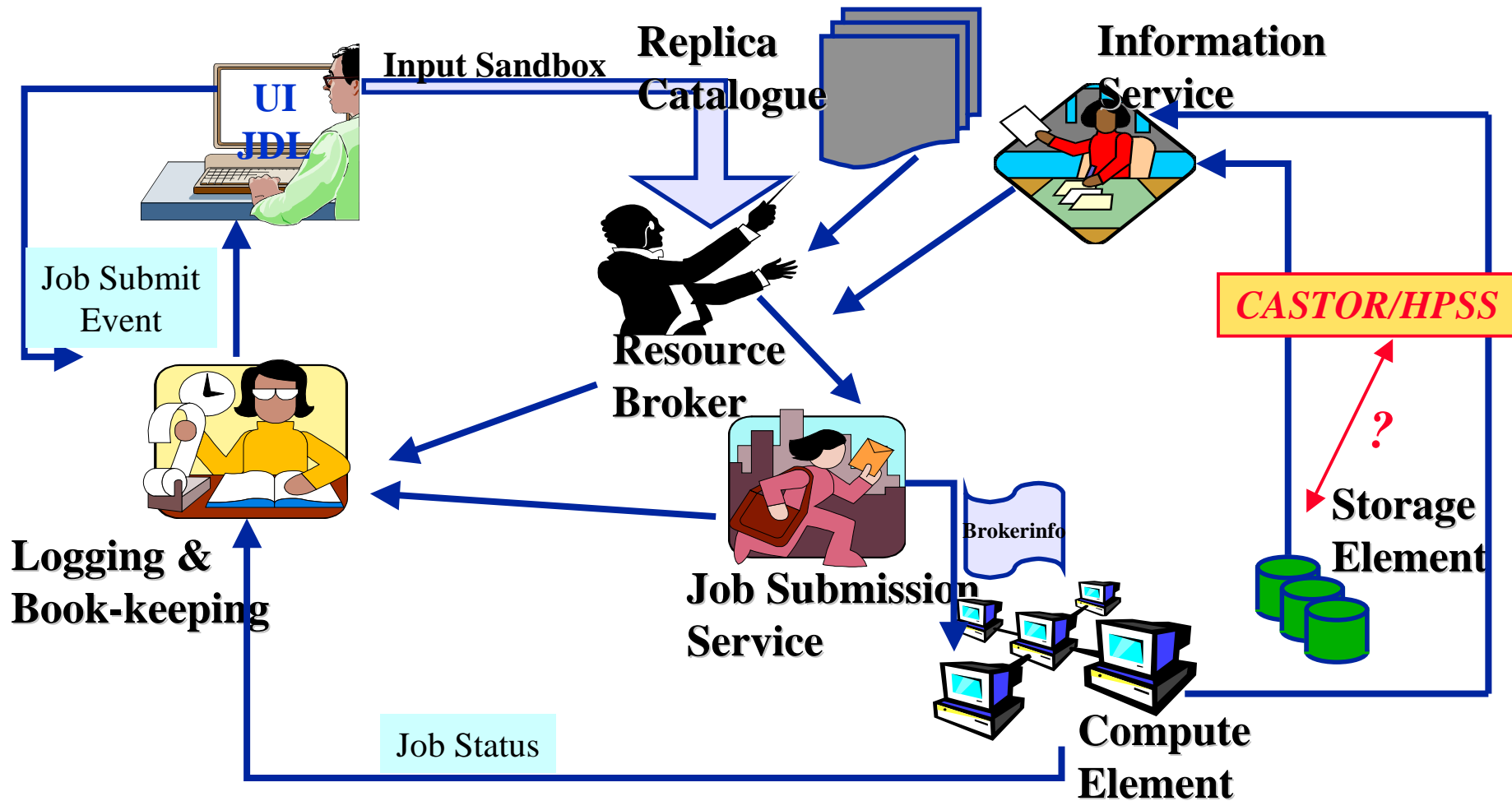
- *permet le monitoring*
- *publie régulièrement les informations renouvelées des sites connectés: par exemple, le nombre de CPUs, les jobs qui tournent, l'espace disque disponible, les VOs autorisés etc...)*

▪ *Replica Manager et Transfert de données*

- *Les outils du Replica Manager et GridFTP permettent des mouvements de données, ie., transfert GridFTP utilisant le mécanisme d'authentification GSI (Globus)*

▪ *Stockage SE*

Les services offerts par EDG (suite ...)



Ce qui a été fait en septembre 2002



- Formation d'une équipe de personnes "Task Force", resp. O. Smirnova
- ATLAS software (release 3.2.1) sous forme de RPM s, distribué, ATLAS EDG Testbed 1.2
- DC1 production script "gridifié", script de soumission de jobs (simulation), préparé et distribué

Input file → simulation dans le détecteur → output files



Un script (JDL) d'application

```
Executable = "/usr/bin/time";
Arguments = "/bin/sh dc1-ds2000.sh 128";
StdOutput = "dc1.002000.simul.00128.hlt.pythia_jet_17.log";
StdError = "dc1.002000.simul.00128.hlt.pythia_jet_17.err";
InputSandbox = {"dc1-ds2000.sh", "rc.conf"};
OutputSandbox =
{"dc1.002000.simul.00128.hlt.pythia_jet_17.log", "dc1.002000.simul.00128.hlt.pyth
ia_jet_17.his", "dc1.002000.simul.00128.hlt.pythia_jet_17.MAG", "dc1.002000.simul.
00128.hlt.pythia_jet_17.AMI", "dc1.002000.simul.00128.hlt.pythia_jet_17.err"};
InputData =
{"LF:evgen/data/002000/dc1.002000.evgen.0007.hlt.pythia_jet_17.root"};
DataAccessProtocol = {"rfio"};
ReplicaCatalog = "ldap://dell04.cnaf.infn.it:9011/lc=Atlas Lc1,rc=ATLAS Testbed1
Replica Catalog,dc=dell04,dc=cnaf,dc=infn,dc=it";
Requirements = other.MaxCPUTime > 86400 &&
Member(other.RunTimeEnvironment, "ATLAS-3.2.1");
Rank = other.FreeCPUs * 1000000 - other.EstimatedTraversalTime;
RetryCount = 0;
```

Ce qui a été fait en septembre 2002 (suite ...)



- 5 fichiers input du Dataset 2000 sont répliqués sur 5 sites, 5 chercheurs × 20 jobs
- Si problèmes: jobs ressoumis

<i>Location</i>	<i>Input partition nr.</i>
CERN	0001, 0002, 0003, 0004, 0005
RAL	0001, 0003
CNAF	0002, 0005
NIKHEF	0003, 0004
CCIN 2P3	0002, 0004

Résultats:

- Succès même si beaucoup de jobs ratés
- Outputs copiés au CERN SE
- 500000 événements simulés, 50 Gb outputs
- Jobs longs > 24 h CPU

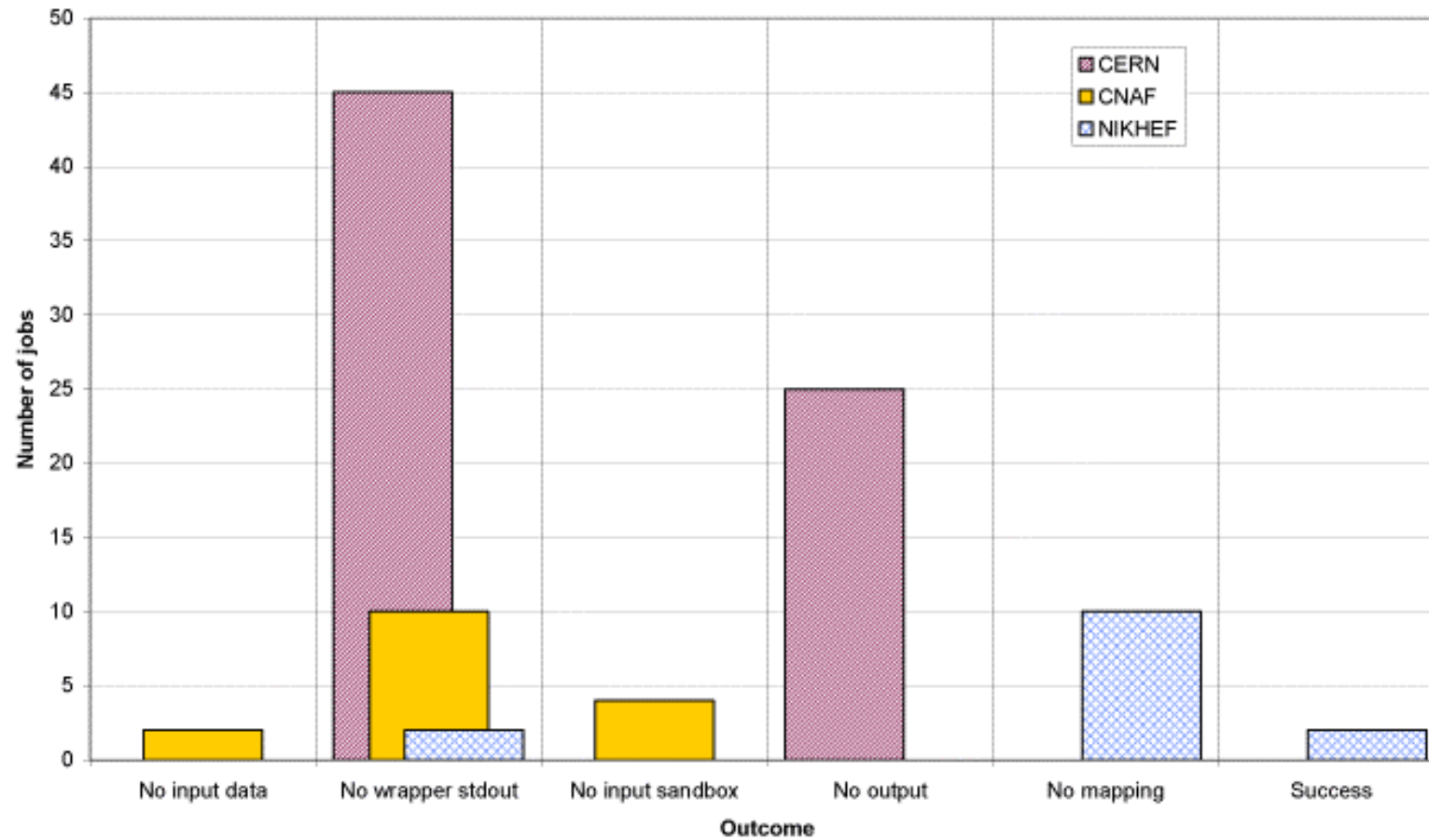


Ce qui a été fait en décembre 2002

- Process (simulation) 5 partitions (fichiers input) du Dataset 2000 @ the EDG Testbed 1.4
- 100 jobs à soumettre, pas de ressoumission
- Output partitions: Tous à enregistrer au CERN SE
- La tâche est pratiquement la même que celle de septembre, sauf que:
 - Pas de ressoumission
 - Un nouveau server pour le Replica Catalog (même site, machine différente)

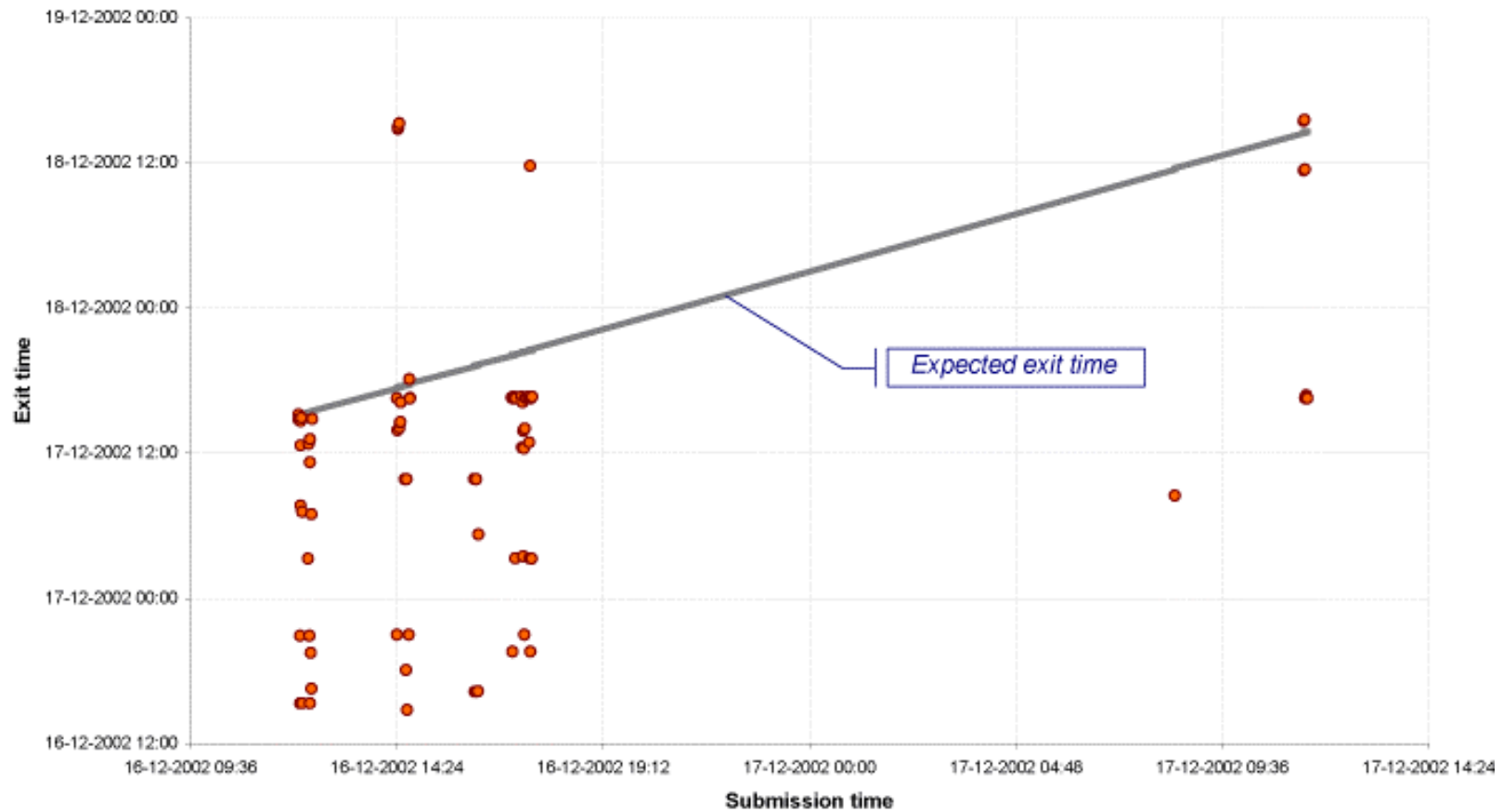
Résultats de décembre 2002

Distribution des jobs et résultats



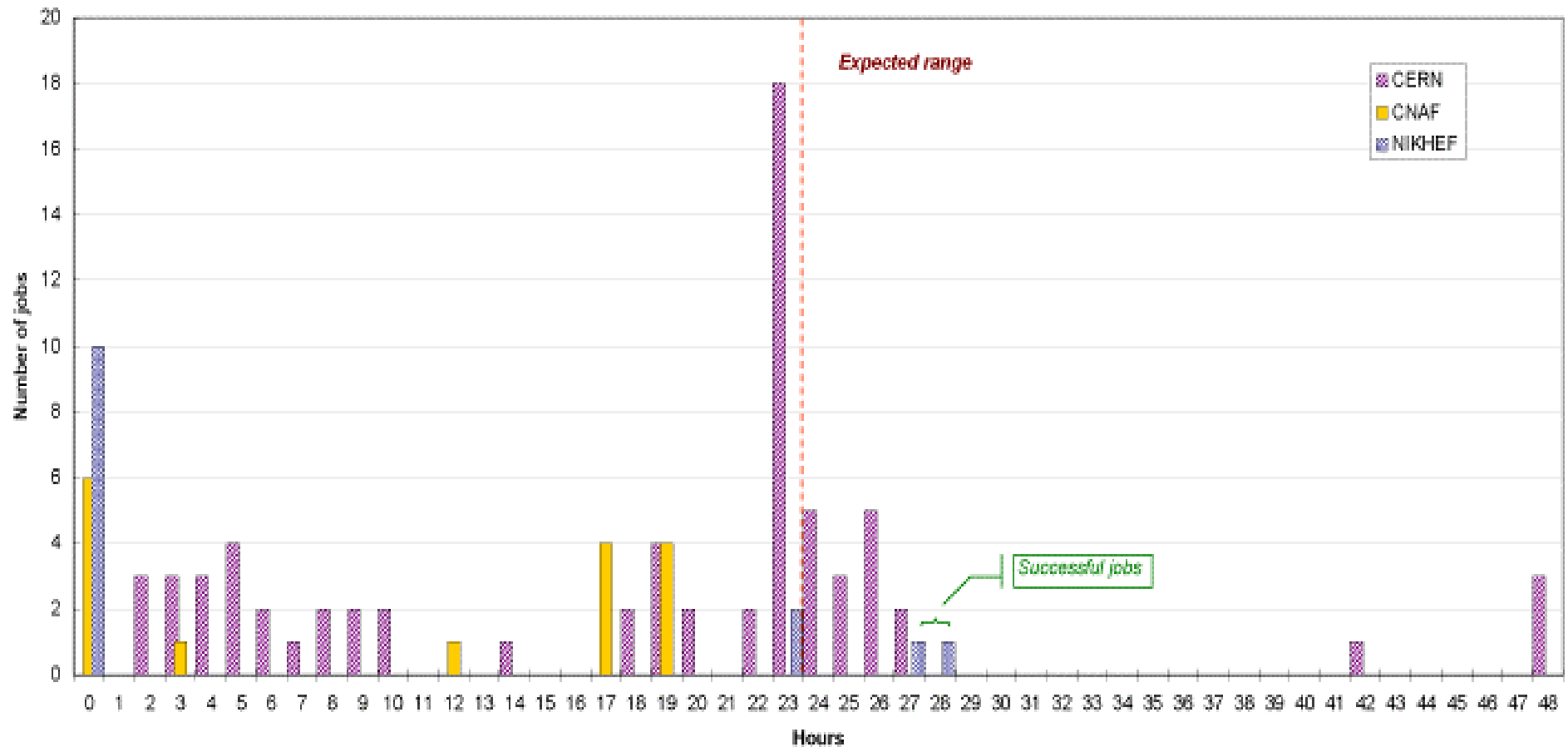
Résultats de décembre 2002

Progression temporelle



Résultats de décembre 2002

Durée des Jobs



Résultats de décembre 2002

Conclusions



- Seulement 2 jobs ont fini complètement
- Problème général avec les longs jobs
 - GASS cache?
 - Globus - PBS interaction?
 - Autre ?
- Problème d'écriture des fichiers output
 - EDG Replica Manager?
 - GridFTP server?
 - Autre ?
- MDS information "fausse" sur certains sites (ETT, free CPU s)
- Interaction avec CASTOR et HPSS ?

Bilan « Atlas » de l'utilisation de EDG



■ Inconvénients

- EDG est encore dans l'état de développement
- non recommandé pour de la production de masse

■ Avantages

- Permet une coopération mondiale facile
- Accès à des ressources multiples et différentes en utilisant le même certificat

■ Les points à améliorer

- Uh oh, everything... Workload management, data management, information system and monitoring, mass storage management, fabric set up. (O. Smirnova)
- ATLAS doit choisir et sélectionner ses bases de données et éventuellement les gridifier

■ Les résultats

- Le software ATLAS fonctionne bien dans l'environnement Grid
- Il est théoriquement possible d'utiliser GRID pour la production de masse
- EDG doit améliorer le « middleware » pour satisfaire à la demande des expériences



Conclusions ...

- Nouvelle release: EDG 1.4.3
- Premiers tests: meilleurs que les précédents
- Pour le Demo EDG EU :
 - 20 partitions (inputs) récupérés de CASTOR au CERN – utilisant rfcpl et globus-url-copy directement de LXPLUS7 vers CERN SE
 - 400 jobs seront soumis (1 semaine avant la demo)
 - La distribution des jobs sera montrée